

VU Research Portal

Important Topics for Fostering Research Integrity by Research Performing and Research Funding Organizations

Labib, Krishma; Roje, Rea; Bouter, Lex; Widdershoven, Guy; Evans, Natalie; Maruši, Ana; Mokkink, Lidwine; Tjeldink, Joeri

published in

Science and Engineering Ethics
2021

DOI (link to publisher)

[10.1007/s11948-021-00322-9](https://doi.org/10.1007/s11948-021-00322-9)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Labib, K., Roje, R., Bouter, L., Widdershoven, G., Evans, N., Maruši, A., Mokkink, L., & Tjeldink, J. (2021). Important Topics for Fostering Research Integrity by Research Performing and Research Funding Organizations: A Delphi Consensus Study. *Science and Engineering Ethics*, 27(4), 1-22. [47].
<https://doi.org/10.1007/s11948-021-00322-9>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

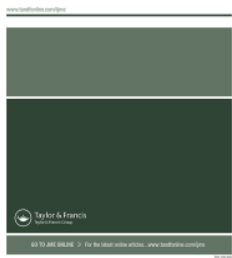
- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl



Reply to the concerns raised by McKenna and Heaney about COSMIN

Lidwine B. Mokkink, Caroline B. Terwee, Lex M. Bouter, Jordi A. Alonso, D. Patrick & Henrica C. W de Vet

To cite this article: Lidwine B. Mokkink, Caroline B. Terwee, Lex M. Bouter, Jordi A. Alonso, D. Patrick & Henrica C. W de Vet (2021) Reply to the concerns raised by McKenna and Heaney about COSMIN, Journal of Medical Economics, 24:1, 857-859, DOI: [10.1080/13696998.2021.1948231](https://doi.org/10.1080/13696998.2021.1948231)

To link to this article: <https://doi.org/10.1080/13696998.2021.1948231>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 12 Jul 2021.



Submit your article to this journal [↗](#)



Article views: 87



View related articles [↗](#)



View Crossmark data [↗](#)

Reply to the concerns raised by McKenna and Heaney about COSMIN

Dear Editor,

We would like to respond to the paper by McKenna and Heaney¹, in which they express their concerns with regard to the COSMIN methodology². We appreciate the attention to our work, and more in general, to the quality of outcome measurement instruments. We very much agree on many of the issues raised. Assessing the quality of measurement instruments is a challenging task which requires knowledge of the construct to be measured, the patient population that is being measured, and the methods of quality assessment of measurement instruments. Therefore, the COSMIN guidelines strongly recommend to conduct systematic reviews on outcome measurement instruments by a multidisciplinary team consisting of experts on all these aspects³. High quality systematic reviews can provide a comprehensive overview of the measurement properties of existing PROMs and supports evidence-based recommendations in the selection of the most suitable PROM available². As these reviews are challenging, complex and time-consuming to conduct, we developed a systematic methodology to allow systematic and transparent processing of all evidence. In addition to the COSMIN methodology for conducting systematic reviews on PROMs², including the COSMIN Risk of Bias checklist for PROMs^{4,5}, we developed the COSMIN Risk of Bias tool for assessing reliability and measurement error for any type of instrument⁶, the COSMIN Study Design checklist⁷ and the COSMIN reporting guideline for studies on measurement properties of PROMs⁸. Note, that each checklist (often accompanied by a peer-reviewed publication, and a user manual) has its own specific purpose.

McKenna and Heaney point to the importance of using conceptual models when developing measurement instruments and evaluating their measurement properties, and suggest that the COSMIN methodology doesn't pay enough attention to defining the construct by grounding it in a theory¹. However, in the COSMIN Risk of Bias checklist, the standards 1 and 2 in Box 1 PROM development, refer to the clarity of the construct, and the use of a conceptual model, respectively⁵. Moreover, in the COSMIN methodology defining the construct is the starting point of a systematic review. Authors of reviews should clearly define the scope of the review, in terms of the construct, the target population, and the context of use⁹. This scope is the reference point for assessing content validity⁹ and for formulating hypotheses for assessing construct validity or responsiveness³. It is up to the reviewers to decide whether they consider a construct to be clearly described, or whether the origin of the construct is clear⁹. However, this is difficult to determine. We agree to

McKenna and Heaney that we reviewers should generally be more strict in this judgement.

A second point these authors raised refers to the use of unidimensional scales, suggesting that within the COSMIN methodology the importance of this is ignored¹. However, COSMIN also addresses that evidence is required on the unidimensionality of (sub)scales based on a reflective model. In the COSMIN methodology^{3,9} each subscale of a multi-dimensional PROM should be considered as a separate measurement instrument. Moreover, internal consistency should be assessed only for unidimensional (sub)scales (Standard 1 Box 4 Internal Consistency), and the assessment of internal consistency (by means of a Cronbach alpha) can only be sufficient when there is proof that the scale is unidimensional (criteria for sufficient internal consistency, see Table 1 of Prinsen et al.²).

McKenna and Heaney clearly explain the advantages of *developing* PROMs using Rasch or IRT analyses¹. Indeed, the CLIQ PROM¹⁰ developed by McKenna and colleagues, was developed based on qualitative studies, in combination with Rasch analyses. We agree that the combination of these two methods is the best method for developing instruments. In the COSMIN methodology, we therefore first consider content validity and explicitly label that as the most important measurement property². Furthermore, when assessing the content validity, the development of the PROM is explicitly taken into account⁵. Standards to assess the quality of the development of a PROM or a study on content validity refer to qualitative methods. The next step is to assess structural validity². Therefore, we developed standards and criteria for studies using factor analyses methods, and IRT or Rasch methods. Nevertheless, when evaluating an instrument, we think IRT or Rasch methods and CTT methods complement each other. Both provide different kinds of information on the quality of an instrument. Moreover, the goal of a systematic review of measurement instruments is often to select the most suitable instrument that is available for use. A systematic review will obviously be restricted to existing instruments and studies, which are still often based on CTT. However, in recent years there is an increase in studies using IRT and Rasch methods. We acknowledge that the COSMIN standards and criteria for these methods should be developed further. We also agree with McKenna and Heaney's call to develop instruments based on IRT or Rasch to achieve scales with interval or ratio level data.

Next, McKenna and Heaney suggest that the COSMIN checklist is based on opinion and not on evidence. The COSMIN checklist was developed using Delphi studies^{5,6,11}. The proposals which were proposed to the experts were

based on existing methodological literature. We invited measurement experts to contribute to the development, and asked them after each proposal to give their arguments and/or provide literature suggestions for their opinion. These arguments and methodological literature were definitively considered when developing the standards and criteria. So, while we agree that opinions were indeed the foundation of the COSMIN methodology, we would like to emphasize that these were opinions of experts who weighted and considered the arguments given for these opinions until a consensus was reached.

A legitimate concern McKenna and Heaney raise is the difference in conclusions drawn by different authors of systematic reviews using the COSMIN methodology¹. The question is whether this is due to the COSMIN methodology or to the expertise and use of COSMIN by these authors. The COSMIN Risk of Bias checklist is not a cookbook. Expertise is required and subjective judgement is sometimes needed. Again, we point to the requirement of having a review team with broad expertise, on both the construct and patient population, but also on the methodology of the type of studies that are included, including both CTT and IRT or Rasch based studies. Furthermore, we facilitate users as much as we can by teaching courses of clinimetrics^{12,13}, and we are currently developing an online course for conducting systematic reviews using the COSMIN methodology. To further increase the quality of systematic reviews of measurement instruments, we recommend users of the COSMIN methodology to be transparent in the methods used, including publishing the search strategy that they used to find studies. We subsequently recommend to publish all results found in the included studies (per measurement property per measurement instrument) in extensive tables, for which we provided examples (see Appendix 7 of the user manual³). This will contribute to the transparency of the conclusions drawn, and allows end users of the review to formulate their own conclusions.

In addition to these important points raised by McKenna and Heaney, we would like to clarify a misunderstanding they seem to have about the COSMIN methodology. The final ratings for the COSMIN Risk of Bias checklist are a worst-score count rating per study per measurement property¹⁴. We did not state that a PROM should have passed 75% of the checklist, as McKenna and Heaney wrongly inferred¹. What we recommend to do, after making this overview of all available evidence, is to categorize the instruments, into (A) PROMs that have potential to be recommended as the most suitable PROM for the construct and population of interest; (B) PROMs that may have the potential to be recommended, but further validation studies are needed; and (C) PROMs that should not be recommended (because of high quality evidence for insufficient measurement properties)².

Assessing the quality of PROMs and other measurement instruments is complex. COSMIN aims to develop a methodology and practical tools to improve the quality of outcome

measurement instruments used in research and clinical practice. We welcome suggestions to further improve the COSMIN materials to contribute to that goal and we invite others to join us in the effort.

Transparency

Declaration of funding

No funding was received to produce this article.

Declaration of financial/other relationships

LBM, CBT, and HCWV receive royalties for the book *Measurement in Medicine* to which we refer in the letter. The royalties are transferred to our university to enable us to do more research.

Acknowledgements

None.

References

- [1] McKenna SP, Heaney A. Setting and maintaining standards for patient-reported outcome measures: Can we rely on the COSMIN checklists? *J Med Econ.* 2021;24(1):502–511.
- [2] Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res.* 2018;27(5):1147–1157.
- [3] Mokkink LB, Vet HC, Prinsen CA, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) - user manual 2018. Available from: www.cosmin.nl
- [4] Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res.* 2018;27(5):1171–1179.
- [5] Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res.* 2018;27(5):1159–1170.
- [6] Mokkink LB, Boers M, van der Vleuten CPM, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med Res Methodol.* 2020;20(1):293.
- [7] Mokkink LB, Prinsen C, Patrick DL, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Version July 2019. 2019 Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf#
- [8] Gagnier JJ, Lai J, Mokkink LB, et al. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res.* 2021. DOI:10.1007/s11136-021-02822-4
- [9] Terwee CB, Prinsen CA, de Vet HCW, et al. COSMIN methodology for assessing the content validity of Patient-Reported Outcome Measures (PROMs). User Manual. 2018. Available from: <https://cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf>
- [10] Wilburn J, McKenna SP, Twiss J, et al. Assessing quality of life in Crohn's disease: development and validation of the Crohn's Life Impact Questionnaire (CLIQ). *Qual Life Res.* 2015;24(9):2279–2288.
- [11] Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* 2010;19(4):539–549.
- [12] de Vet HCW, Terwee CB, Mokkink LB. *Clinimetrics: assessing measurement properties of health measurement instruments.* Available

from: <https://www.epidm.nl/en/courses/clinimetrics-assessing-measurement-properties-health-measurement-instruments/>.

- [13] de Vet HC, Terwee CB, Mokkink L, et al. *Measurement in medicine*. Cambridge: Cambridge University Press; 2011.
- [14] Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012;21(4):651–657.

Lidwine B. Mokkink, Caroline B. Terwee and
Henrica C. W. de Vet

*Department of Epidemiology and Data Science, Amsterdam
Public Health Research Institute, Amsterdam UMC, Vrije
Universiteit Amsterdam, Amsterdam, The Netherlands*

Lex M. Bouter

*Department of Epidemiology and Data Science, Amsterdam
Public Health Research Institute, Amsterdam UMC, Vrije
Universiteit Amsterdam, Amsterdam, The Netherlands
Department of Philosophy, Faculty of Humanities, Vrije
Universiteit Amsterdam, Amsterdam, The Netherlands*

Jordi A. Alonso

*Health Services Research Group, IMIM-Hospital del Mar Medical
Research Institute; CIBER en Epidemiología y Salud Pública
(CIBERESP); Pompeu Fabra University, Barcelona, Spain*

D. Patrick

*Department of Health Services, University of Washington,
Seattle, WA, USA*

✉ w.mokkink@amsterdamumc.nl

Received 18 May 2021; accepted 23 June 2021

© 2021 The Author(s). Published by Informa UK Limited, trading as
Taylor & Francis Group.

This is an Open Access article distributed under the terms of the
Creative Commons Attribution-NonCommercial-NoDerivatives License
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-
commercial re-use, distribution, and reproduction in any medium,
provided the original work is properly cited, and is not altered,
transformed, or built upon in any way.